

New Approaches to Structure-Based Discovery of Dengue Protease Inhibitors

S. M. Tomlinson[#], R. D. Malmstrom[#] and S. J. Watowich*

Department of Biochemistry and Molecular Biology, and Sealy Center for Structural Biology and Molecular Biophysics, University of Texas Medical Branch, Galveston, Texas 77555, USA

Abstract. Dengue virus (DENV), a member of the family *Flaviviridae*, presents a tremendous threat to global health since an estimated 2.5 billion people worldwide are at risk for epidemic transmission. DENV infections are primarily restricted to sub-tropical and tropical regions; however, there is concern that the virus will spread into new regions including the United States [1]. There are no approved antiviral drugs or vaccines to combat dengue infection, although DENV vaccines have entered Phase 3 clinical trials. Drug discovery and development efforts against DENV and other viral pathogens must overcome specificity, efficacy, safety, and resistance challenges before the shortage of licensed drugs to treat viral infections can be relieved. Current drug discovery methods are largely inefficient and thus relatively ineffective at tackling the growing threat to public health presented by emerging and reemerging viral pathogens. This review discusses current and newly implemented structure-based computational efforts to discover antivirals that target the DENV NS3 protease, although it is clear that these computational tools can be applied to most disease targets.

SIGNIFICANCE

Dengue virus (DENV) is one of the most important global pathogens and may represent a global pandemic [2]. Approximately 2.5 billion people worldwide at risk for dengue infection [3]. Between 50-100 million dengue infections are estimated to occur each year, with upwards of 1.5 million infected individuals presenting with clinical symptoms and ~500,000 infections progressing to the serious disease states of dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS). Each year, dengue infections result in ~25 000 deaths, primarily in children.

TRANSMISSION

After malaria, DENV is the most common mosquito-borne pathogen that infects humans. Mosquitoes of the *Aedes* genus transmit DENV between infected human hosts, with the main vector *Aedes aegypti* well adapted to urban environments. *Aedes albopictus* and *Aedes polynesiensis* have also been implicated in dengue outbreaks [4]. Resurgence of dengue disease in the 1970's in Central and South America occurred after cessation of mosquito control measures designed to reduce dengue epidemics. Population growth, urbanization, increased air travel, and climate change may also be factors in the global reemergence of dengue disease.

INCIDENCE AND PREVALENCE

Reports of dengue epidemics date back more than 200 years ago in Africa, Asia, and North America [5]. Dengue virus has been reported in more than 100 countries [6], predominantly in tropical and subtropical regions (Fig. 1).

Dengue outbreaks have been reported throughout these regions with increasing frequency and disease severity (Fig. 2), although it is unknown what fraction of the recent increases in dengue disease are due to improved surveillance and diagnosis. In light of the recent outbreak of the related West Nile virus in North America [7], and the fact that the mosquito vectors that transmit dengue appear to be expanding their range, the U.S. National Institutes of Health has expressed concern that dengue disease represents a potential threat to the United States [1].

DENGUE DISEASE

Dengue viruses can be classified into four distinct serotypes (dengue 1-4) based on antigenic properties or sequence similarity (Fig. 3). Infection by one serotype does not protect from, but may actually enhance, subsequent infection by any of the other three serotypes [8]. Infections may be asymptomatic or may present with varying disease symptoms and severity. The incubation period for dengue ranges from 3-8 days. Dengue fever (DF), the most common form of the disease, is most often characterized by typical flu-like symptoms, including fever, fatigue, rash, severe headache, nausea, vomiting, diarrhea, myalgia, arthralgia, retro-orbital pain, and minor hemorrhagic manifestations such as epistaxis, petechiae, and gingival bleeding [9]. Early symptoms of DHF, which most often occurs in children, are the same as those for dengue fever. Hemorrhagic symptoms, however, are much more severe in DHF than in typical DF with the primary pathophysiologic contributor being plasma leakage resulting in hemoconcentration, thrombocytopenia, pleural effusions, hypoproteinemia, and/or hypoalbuminemia [10]. While the first reported DHF outbreak occurred in Manila in the 1950's, epidemics have become increasingly more frequent. It is believed that repeat infection by a different serotype of the virus increases the probability of developing the more severe form of the disease [8], hence people living in endemic areas where more than one serotype

*Address correspondence to this author at the Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, TX 77555-0647; Tel: 409-747-4749; Fax: 409-747-4745; E-mail: watowich@xray.utmb.edu

[#]These authors contributed equally.

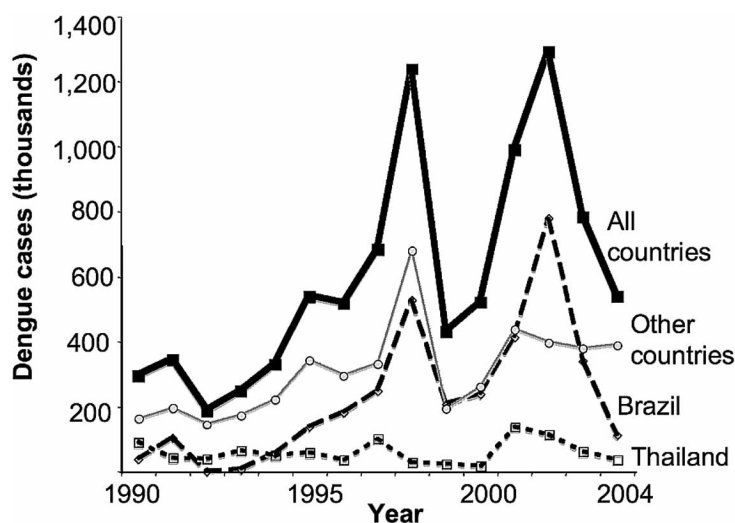


Fig. (1). Annual reported clinical cases of DENV infection between 1990 and 2004. Individual country data was plotted for Thailand as representative of Southeast Asia (open squares) and Brazil since it contains the greatest number of reported cases in the Americas (open diamonds). Annual worldwide cases of DENV infection between 1990 and 2004 were represented by filled squares. The shaded circles and line labeled “Other countries” represented case incidence data from all countries other than Thailand and Brazil. Data was obtained from WHO DengueNet (www.who.int/globalatlas). Data between 2005–2008 were largely incomplete and not used in this figure.

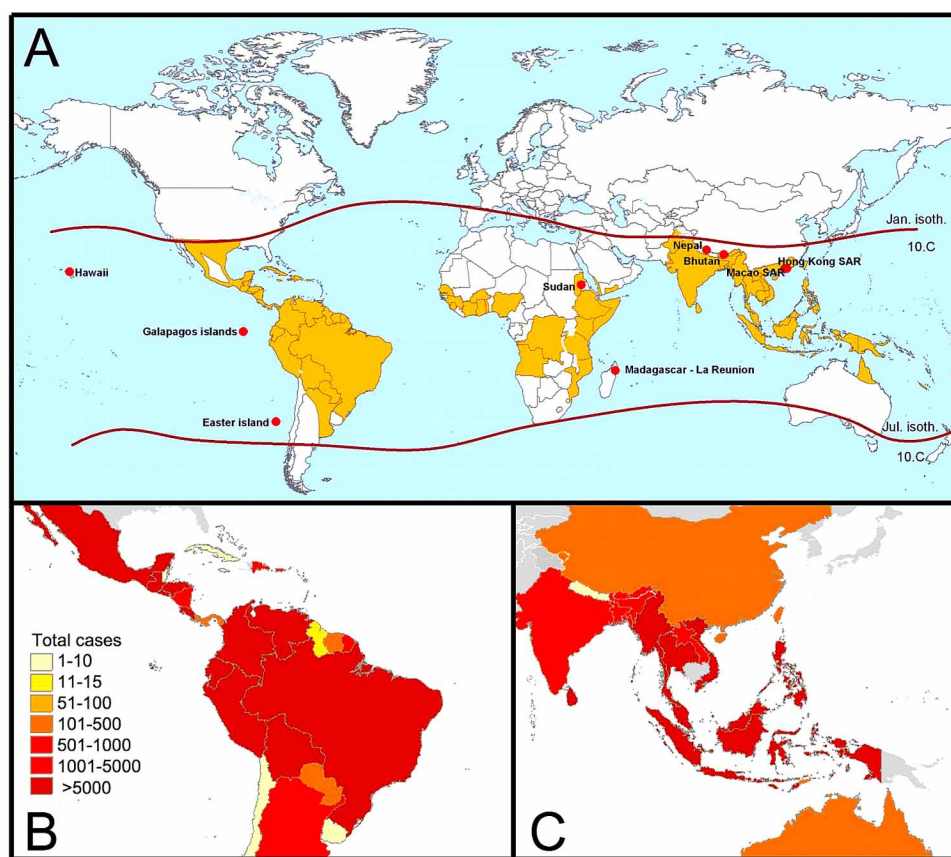


Fig. (2). (A) Approximate distribution of DENV cases in 2008. Cases were largely confined to subtropical and tropical regions of the world. The map was adapted from WHO DengueNet (www.who.int/globalatlas). (B/C) Maps showing the number of reported cases of DENV infection within individual countries of (B) South America and (C) Southeast Asia. Figures were adapted from maps generated from WHO DengueNet (www.who.int/globalatlas) using data for 2004 (last year with complete statistics). Grey colors indicate no reported data.

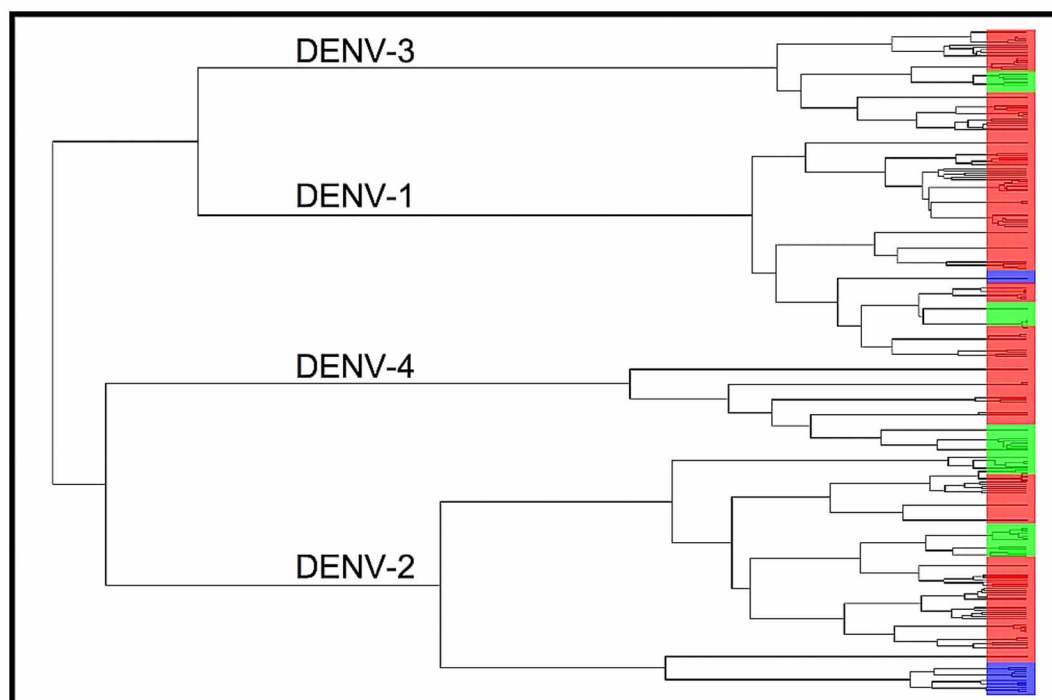


Fig. (3). Phylogenetic tree showing separation of dengue viruses into 4 distinct clades based on sequence similarity of the NS3 protein. Data from 859 sequences isolated between 1944-2008 were retrieved from the National Institutes of Health NCBI Virus Variation database (www.ncbi.nlm.nih.gov/genomes/VirusVariation) and the phylogenetic tree produced using complete linkage further neighbor clustering and Hamming distances. The tree branches largely clustered by geographic region (Asia/Oceania, red; South America, green; Africa, blue) within each clade.

are circulating may be at greater risk for developing DHF. Depending on the severity of the hemorrhagic symptoms, shock may occur resulting in dengue shock syndrome (DSS). Mortality is primarily associated with the more severe disease cases.

TREATMENT

Treatment for dengue disease consists primarily of intensive supportive therapy. Maintaining circulating blood volume is the most important feature in the management of the disease. While healthcare improvements over the last several years in many of the affected regions have resulted in a decrease in the mortality rate associated with the disease, many regions still suffer from inadequate healthcare facilities and inability to treat disease symptoms. The proteins required for the fitness of the virus provide several potential targets against which to develop antiviral drugs. Although several anti-dengue therapies are in clinical studies [11-13], there are currently no approved vaccines to prevent dengue and no antiviral drugs to treat the disease.

FAMILY FLAVIVIRIDAE

Dengue virus belongs to the family *Flaviviridae* that includes a number of important human pathogens such as yellow fever, West Nile, Japanese encephalitis, and hepatitis C viruses. The family contains more than 70 members organized into three genera including the flaviviruses (arthropod-

borne viruses), the pestiviruses, and the hepaciviruses, which include the blood-borne hepatitis C viruses.

MOLECULAR ORGANIZATION

The viruses within the flavivirus genus have similar molecular features. They are enveloped RNA viruses with a positive single strand ~11kb genome. The genome contains a type 1 cap on its 5' end, but the 3' end has no poly-A tail. The genome encodes 3 structural proteins at the 5' end, including the capsid (C), membrane precursor (prM; also termed preM in some literature), and envelope (E) proteins, and 7 nonstructural proteins at the 3' end (Fig. 5). Viruses are icosahedral particles of ~50 nm diameter with a host-derived lipid bilayer encasing a nucleocapsid of viral RNA complexed with 240 copies of the capsid protein.

REPLICATION CYCLE

The detailed functional interactions that occur during viral replication are largely unknown, although a general understanding is emerging as to how dengue proteins regulate the virus replication cycle (Fig. 4). Virus particles enter the cell by receptor-mediated endocytosis. Upon acidification of the endocytic vesicle, the nucleocapsid enters into the cytoplasm where the virus genome is released. The genome is translated into a single polypeptide, and undergoes co-translational and post-translational processing by host and viral proteases to produce the individual proteins required for viral replication and packaging. Replication takes place on

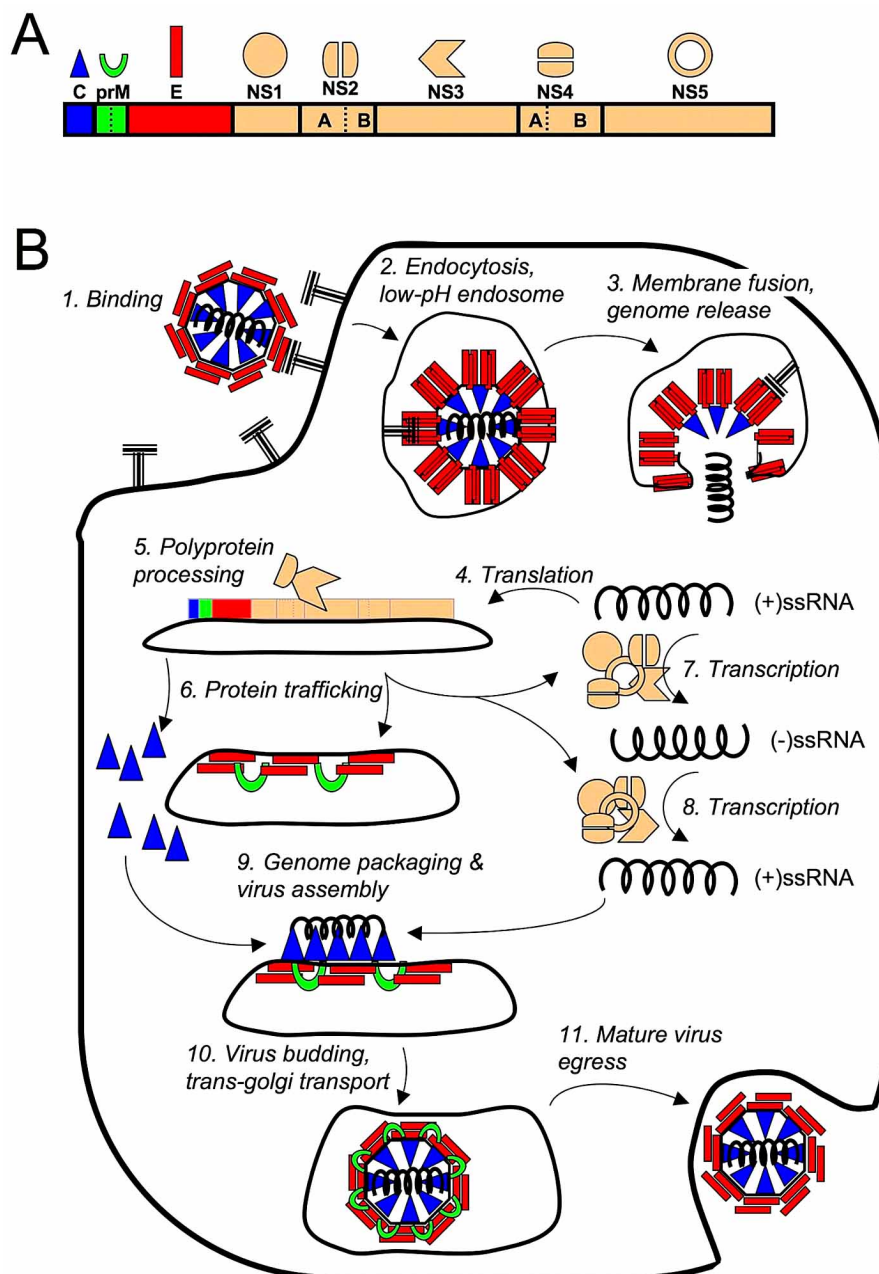


Fig. (4). Schematic diagram of dengue proteins likely involved in the various phases of the dengue virus replication cycle. (A) Diagram of the dengue polyprotein, with cartoon representations for the structural (C, prM, E) and non-structural (NS1, NS2, NS3, NS4, NS5) proteins involved in virus replication. (B) Overview of the dengue virus replication cycle, highlighting virus proteins that likely regulate each stage of replication. The detailed composition of the viral replication complex is largely unknown, but is believed to involve many (if not all) of the non-structural proteins.

intercellular membranes, and assembly occurs on the endoplasmic reticulum (ER) membrane. Newly assembled virus particles are transported through the trans-golgi network and are released from the cell by exocytosis [14]. Since the viral proteins are crucial to the various stages of entry, replication, assembly, and egress (Fig. 4) they serve as potential targets for the development of novel antiviral drugs. High-resolution atomic structures have been deter-

mined for many of the potential target proteins of dengue virus (Fig. 5, Table 1).

DENGUE VIRUS STRUCTURAL PROTEINS AS POTENTIAL DRUG TARGETS

Virus particles enter the host cell by receptor-mediated endocytosis. As the endocytic vesicles acidify, the viral E

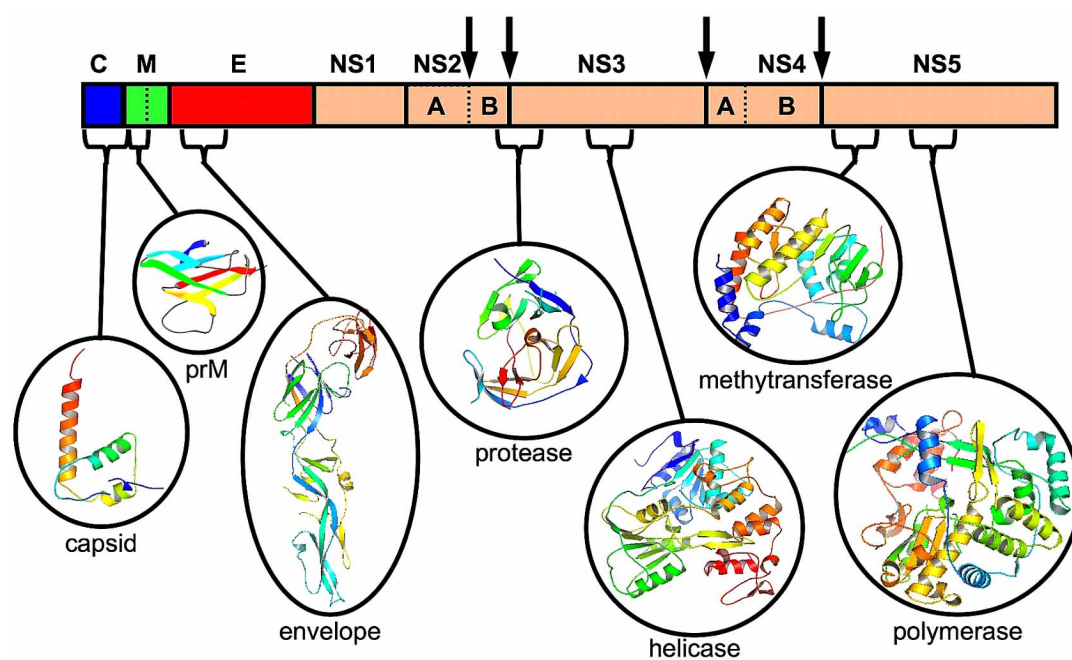


Fig. (5). Structural elements of the DENV polyprotein. Structures were shown for individual dengue proteins determined to atomic resolution suitable for structure-based drug discovery and design. Vertical arrows indicate major polyprotein cleavage sites recognized by the dengue NS2B-NS3 protease.

Table 1. Representative Structures Available in RCSB for DENV Target Proteins

Target	PDB	Method	Resolution	Comments	Reference
E and M	1p58	cryo-EM	9.5 Å	Mature virion	[138]
E	1thd	cryo-EM, PA*	9.5 Å	Mature virus particle	[139]
E	1tge	cryo-EM	12.5 Å	Immature virus particle	[139]
E	1tg8	X-ray	2.61 Å		[139]
E	1ok8	X-ray	2.0 Å	Post-fusion conformation	[140]
E	1uzg	X-ray	3.5 Å	Dengue 3	[141]
E	2b6b	cryo-EM		Complexed with crd of DC-SIGN	[142]
E	1oke	X-ray	2.4 Å	Complex with N-octyl-β-D-glucoside	[18]
E	1oan	X-ray	2.75 Å	Ligand-binding pocket / apo-protein	[18]
E	1k4r	cryo-EM	24 Å	First cryo-EM structure	[143]
E	2r29	X-ray	3 Å	Binding with neutralizing antibody	[34]
E	2r6p	cryo-EM, PA*	24 Å	E protein and Fab 1A1D-2 into the cryo-EM structure of Fab/Dengue complex	[34]
E	3c6r	cryo-EM	25 Å	Low pH immature virion	[144]
E	2r69	X-ray	3.8 Å	Fab 1A1D-2 / Domain III	[34]
prM / E	3c6d	cryo-EM	12.5 Å	Immature virion	[145]
E	2jsf	NMR		Domain III	[33]

(Table 1) Contd.....

Target	PDB	Method	Resolution	Comments	Reference
E	2h0p	NMR		Dengue 4 Domain III	[146]
prM	1n6g	cryo-EM	16 Å	Immature dengue-2 PrM	[138]
prM / E	3c5x	X-ray	2.2 Å	PrM / E heterodimer at low pH	[145]
prM / E	3c6e	X-ray	2.6 Å	PrM / E heterodimer at neutral pH	[145]
Capsid	1r6r	NMR		21 structures	[26]
NS3 protease	1bef	X-ray	2.1 Å	Apo structure	[63]
NS3 protease	1df9	X-ray	2.1 Å	Complex with Mung-Bean Bowman-Birk inhibitor	[64]
NS2B-NS3 protease/cofactor	2fom	X-ray	1.5 Å	Active protease	[68]
NS3 protease	2qid	X-ray	2.1 Å	Complex with Mung-Bean Bowman-Birk inhibitor	Murthy 2007
NS3 protease/helicase	2vbc	X-ray	3.15 Å		[40]
NS3 helicase/NTPase	2bhr	X-ray	2.8 Å		[39]
NS3 helicase/NTPase	2bmf	X-ray	2.41 Å		[39]
Methyltransferase	2p3o	X-ray	2.76 Å	Complex with 7MeGpppA and S-adenosyl-L-homocysteine	[46]
Methyltransferase	2p3l	X-ray	2.20 Å	Complex with GpppA and S-adenosyl-L-Homocysteine	[46]
Methyltransferase	2p3q	X-ray	2.75 Å	Complex with GpppG and S-adenosyl-L-homocysteine	[46]
Methyltransferase	2p40	X-ray	2.70 Å	Complex with 7MeGpppG	[46]
Methyltransferase	2p41	X-ray	1.80 Å	Complex with 7-MeGpppG2'OMe and S-adenosyl-L-homocysteine	[46]
Methyltransferase	1r6a	X-ray	2.60 Å	Complex with s-adenosyl homocysteine and ribavirin 5' triphosphate	[44]
Methyltransferase	2p1d	X-ray	2.90 Å	Complex with GTP and S-adenosyl-L-homocysteine	[45]
Methyltransferase	119k	X-ray	2.40 Å		[45]
Polymerase	2j7u	X-ray	1.85 Å	Apo-structure	[48]
Polymerase	2j7w	X-ray	2.60 Å	Complex with 3'DGTP	[48]

*PA - pseudo-atomic modeling

protein trimerizes and mediates virus and cell membrane fusion [15]. Following membrane fusion, the nucleocapsid enters the cytoplasm and uncoats to release the viral genome [16]. Numerous studies have investigated approaches to disrupt virus infection by interfering with viral entry. Key proteins in these interactions are the E (envelope), prM (premembrane), and C (capsid) structural proteins. Potential entry-associated inhibitors that have been evaluated include small molecules [17-19] and peptide inhibitors [20].

The capsid (C) protein is necessary for packaging and release of mature viral particles from the cell. The capsid C-terminal hydrophobic region contains a signal sequence for anchoring the protein into the ER membrane and partitioning the prM protein to the membrane [21]. The central hydro-

phobic portion of the capsid protein promotes association of the capsid protein and ER membrane [22], and has been shown in related flaviviruses to be important for efficient virus assembly [23]. The highly basic regions of the capsid protein have been implicated in interactions with viral RNA [24, 25]. An NMR structure of the capsid protein [26] provided supporting structural evidence that the protein's highly basic region could interact with viral RNA and its hydrophobic region could interact with lipid molecules.

The prM protein is the precursor of the M protein. The largely hydrophilic N-terminal portion of the protein codes for the glycosylated "pr" segment of the prM protein; this protein provides resistance to low pH environments [27] and protects the envelope protein from premature fusion during

transit through the acidic environment of the trans-golgi network. Cleavage of the pr peptide from prM generates the M protein. This cleavage is mediated by the cellular protease furin and is required for virus maturation [28] and infection in cell culture [29]. Insights into the function of this protein have been aided by pH-dependent cryo-EM and X-ray structural studies (Table 1).

The envelope (E) protein consists of three domains. Domain I, the N-terminal portion of the protein, is centrally located in the 3D structure. Domain II contains the sequence for a class II fusion peptide that is involved in viral fusion with the endosomal membrane and release of the genome into the host cytosol. Based on solution [30] and X-ray crystallographic studies [31] of tick-borne flavivirus E proteins, domain III was reported to be the putative receptor-binding domain for virus attachment to the host cell. In addition, this domain was shown to be involved with antibody neutralization [32-35]. Prior to endocytosis, the E protein is anchored to the viral membrane. Upon acidification of the endosome, the E protein undergoes conformational changes that are believed to induce fusion of the virus and cell membranes and promote release of the virus nucleocapsid into the cytoplasm of the infected cell. The E protein conformation changes between the immature, mature, and fusion-activated states of the virus [36]. Many of the dengue protein structures reported in the literature are for the E protein (Table 1), reflecting the range and importance of E protein conformational changes required for virus fitness.

DENGUE VIRUS NONSTRUCTURAL PROTEINS AS POTENTIAL DRUG TARGETS

The nonstructural (NS) enzymes of the replication complex include the NS3 protease and with its NS2B cofactor, the NS3 helicase/nucleoside triphosphatase (NTPase)/RNA 5' triphosphatase (RTPase), and the NS5 methyltransferase/RNA-dependent RNA polymerase. These proteins serve as potential inhibitory targets for antiviral agents since they are required for virus replication.

The multi-functional C-terminal domain of NS3 encodes NTPase, helicase, and RTPase activities. NTP hydrolysis is thought to provide the chemical energy required for helicase activity [37]. The likely function of the helicase is to unwind double-stranded RNA during viral replication; it is required for viral replication [38]. These functional roles were supported by recent flavivirus helicase/NTPase crystal structures [39]. A recently solved crystal structure of linked NS3 protease/helicase domains suggested the protease domain enhanced binding of RNA to the NS3 helicase [40]. The RTPase is thought to modify the 5' end of the viral RNA to prepare it for cap addition [41]. Mutational studies have revealed a region on the full length NS3 protein required for optimal activity of all three enzymatic functions [42], although little progress has been made identifying inhibitors of DENV helicase, NTPase, and RTPase activities.

The NS5 protein contains methyltransferase (MTase) and RNA-dependent RNA polymerase (RdRp) activity. The N-terminal portion of NS5 encodes the MTase and contains guanlyltransferase activity. Its function is methylation and capping of viral RNA; it is essential for viral replication

[43]. A number of different crystal structures of the methyltransferase [44-46] have helped provide structure-function insights. MTase crystal structures were used for computational virtual screening and identified a small molecule inhibitor with $IC_{50} \sim 60 \mu M$. Ribavirin, which is used in combination with interferon to treat hepatitis C virus (a distantly related hepacivirus) infection, functions by inhibiting MTase activity [44]. The NS5 C-terminal domain contains the RdRp activity. Biochemical high-throughput assays have been developed to test the efficacy of small molecule RdRp inhibitors [47]. Two flavivirus polymerase X-ray crystal structures were recently solved [48] and have provided structure-function insights and suggestions for RdRp structure-based inhibitor design [49]. Polymerase inhibitors are reportedly under active investigation by several pharmaceutical companies (Bloomberg, 23 June 2008), although to our knowledge no studies have entered the peer-reviewed literature.

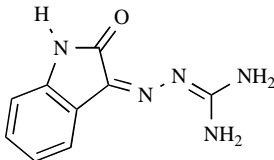
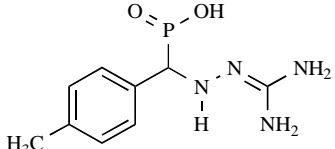
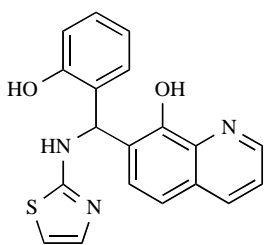
DENGUE VIRUS NS3 PROTEASE AS A POTENTIAL DRUG TARGET

The N-terminal 180 residues of NS3 encode a trypsin-like serine protease [50]. Required for NS3 protease activity is a ~40 residue hydrophilic domain from NS2B [51]. The dengue NS2B-NS3 protease has a preference for dibasic substrates and cleaves the newly translated dengue polyprotein at the NS2A-NS2B, NS2B-NS3, NS3-NS4A, NS4B-NS5 polyprotein cleavage sites as well as internal sites within C, NS2A, NS3 and NS4A proteins [52-55]. The NS3 protease is a primary target for development of dengue antiviral drugs since the NS2-NS3B protease is required for virus replication [56] and protease inhibitors have a successful history as being developed into antiviral drugs [57, 58].

Biochemical assays with active recombinant protease (constructed from 40 residues from the NS2B cofactor linked to the 180 residues N-terminal residues of NS3) and tripeptide fluorogenic [51] or chromogenic [59] substrates lead to quantitative studies of peptide-based protease inhibitors [59, 60]. Preliminary functional group substitution studies identified boronic acid and trifluoromethyl ketone derivatives with nanomolar affinities and an aldehyde derivative with a low micromolar affinity [61] (Table 2). Structure-activity relationships with aldehyde-based peptide inhibitors highlighted the relative importance of the P2 side chain to substrate binding [62].

X-ray crystal structures of the dengue NS3 protease (PDB identifier 1BEF) [63] and of the NS3 protease complexed with the mung-bean Bowman-Birk inhibitor (PDB identifier 1DF9) [64] provided an initial structural basis to understand NS3 protease function (Fig. 6A, C). Although these structures did not represent the active form of the enzyme, they were successfully used in structure-based calculations to identify small molecule protease inhibitors (Table 2) [65, 66]. Recently, the structure of active dengue NS2B-NS3 protease was determined (Fig. 6B, D) [67, 68]. Although the protease catalytic site residues (His51, Asp75, Ser131) were arranged similarly in the NS3 and NS2B-NS3 crystal structures, numerous large conformational differences were evident (Fig. 6). For instance,

Table 2. Representative Dengue NS2B-NS3 Protease Inhibitors

Structure	Ki (μM)	References
	23	[65]
	14	[65]
	30	Mueller 2008
Ac-FAAGRR- α keto-SL-CONH ₂	47	[59]
Ac-FAAGRR-CHO	16	[59]
Ac-FAAGRK-CONH ₂	26	[60]
Ac-EVKKQR-CONH ₂	67	[60]
Ac-RTSKKR-CONH ₂	12	[60]
Ac-TTSTRR-CONH ₂	46	[60]
Ac-KKR-CONH ₂	22	[60]
BZ-NLE-KRR-H	6	[61]
BZ-NLE-KRR-B(OH) ₂	0.04	[61]
BZ-NLE-KRR-CF ₃	0.9	[61]
BZ-NLE-KRR-Thiazole	43	[61]

overlaying the catalytic regions of the two structures resulted in positional differences of 14 Å and 35 Å for Leu31 and Asn119, respectively (Fig. 6). Of relevance to dengue protease inhibitor design are the large differences in the substrate-binding region between the two structures (Fig. 6). The S1 site within the NS3 substrate-binding region formed a deep pocket that could accommodate long positively charged P1 sidechains of the substrate (Fig. 6C). However, in the NS2b-NS3 structure, the S1 site forms only a shallow depression (Fig. 6D). Structure-based drug discovery approaches must consider the differences between the NS3 and NS2B-NS3 structures since small molecules may interact differently with the active sites of NS3 and NS2B-NS3 (Fig. 7) [69].

OTHER STRATEGIES FOR DENGUE ANTIVIRAL DRUG DISCOVERY

Structure-based approaches, although the focus of this review, are not the only methods available to discover dengue antiviral compounds. Alternative approaches include modulating the host immune response, high-throughput screening (HTS) using virus replication cell-based assays, or HTS specifically targeting viral morphogenesis [70], the 3' UTR [71], virus absorption [72], or assembly and maturation [73]. A challenge for inhibitors discovered with virus replication cell-based HTS is determining the mechanism of

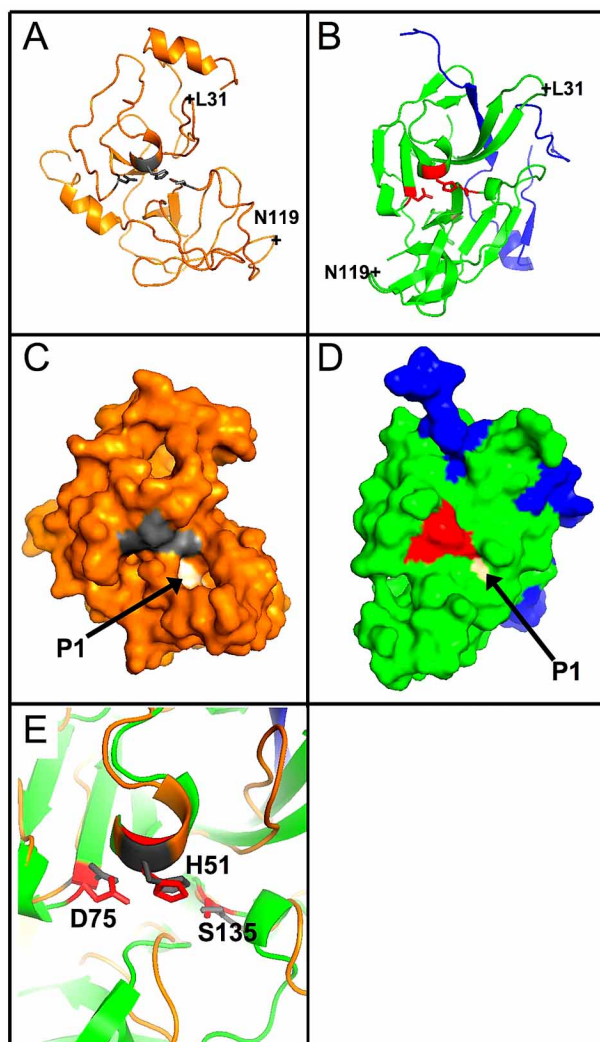


Fig. (6). Comparison of the dengue NS3 protease [63] and the dengue NS2B-NS3 protease [68]. Ribbon diagram representation of (A) NS3 and (B) NS2B-NS3 proteins. The conformation changes between the two structures shifts leucine 31 (L31) and asparagine 119 (N119) by 14 Å and 35 Å, respectively. Molecular surface diagrams for (C) NS3 and (D) NS2B-NS3. Arrows point to the substrate binding P1 pocket to highlight the differences in the structure of the binding pocket in the two structures. (E) Superposition of NS3 and NS2B-NS3 catalytic sites demonstrated that this region was similarly structured in the two proteins. Catalytic residues His51, Asp75 and Ser131 were colored grey and red in the NS3 and NS2B-NS3 structures, respectively.

inhibition. Testing natural products identified several lead compounds that inhibit DENV replication in cell culture [74-77], with the components of fingerroot (*Boesenbergia rotunda*) reported to inhibit the DENV protease with a μM inhibition constant [74]. A very general strategy utilizes compounds identified from other viral studies and tests them for inhibitory activity against dengue replication. There have been few discovery successes utilizing these various strategies, and to our knowledge no leads have progressed to clinical trials.

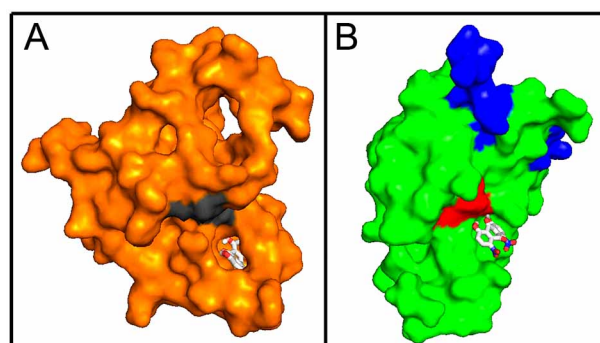


Fig. (7). Comparison of a previously identified WNV inhibitor [69] docked to the molecular surface of (A) DENV NS3 and (B) NS2B-NS3. The NS3 protein molecular surface was colored orange. Molecular surface corresponding to the NS2B and NS3 protein chains of NS2B-NS3 were colored blue and green, respectively. Catalytic residues His51, Asp75 and Ser131 were colored grey and red in the NS3 and NS2B-NS3 structures, respectively.

Viral proteases are a primary target for antiviral discovery since they are common to most viruses and generally important for efficient replication [57]. Although HIV protease inhibitors provide proof-of-principle for the development of antiviral protease inhibitors [57, 58], human rhinovirus [78] and hepatitis C virus [79] protease inhibitors have also entered clinical trials. In spite of these successes, very few antivirals have been approved for human use.

The remainder of this paper focuses on recent improvements to structure-based computational approaches for drug discovery, with an emphasis on targeting the dengue protease. As additional virus protein structures become available, improved computational methods will be needed to efficiently translate structural information to potential antiviral drugs. Moreover, these improved computational approaches and associated high-performance computer resources need to become increasingly available to biomedical researchers.

OVERVIEW OF COMPUTATIONAL METHODS FOR DRUG DISCOVERY

Detailed structural information is becoming increasingly available for the dengue virus proteins (Fig. 5). Since these proteins are required for virus infectivity and replication, structure-based computational approaches [80-82] offer an attractive strategy for the discovery and optimization of dengue antiviral drugs. Moreover, new computational approaches promise to improve the effectiveness of current structure-based calculations.

Virtual screening is a powerful structure-based drug discovery approach that systematically examines each molecule within a chemical compound databases to identify those molecules that maximize a desired pharmacological profile (e.g., binding energy to an enzyme active site). The necessary components of a virtual screening calculation include a database or library of defined chemical compounds (or pharmacophores), a target macromolecule of known three-dimensional structure, a defined region on the macromolecule to restrict the chemical space being exa-

mined, and a docking program that exhaustively examines conformational space of each molecule in the database and assigns each conformation a quantitative score based on a specified property (e.g., binding).

COMPOUND DATABASES

These are virtual collections of defined chemical compounds or pharmacophores that contain hundreds to millions of entries. Some examples include the noncommercial ZINC database [83] with ~8 million small molecules represented as 3-dimensional chemical and electronic structures calculated from chemical bonding patterns, the Cambridge Structural Database [84] which contains 3-dimensional chemical structures determined by X-ray crystallography, PubChem (pubchem.ncbi.nlm.nih.gov) which is a collection of publicly available databases cross-referenced with biological activities and literature citations, and many chemical supply companies which have compiled chemical catalogs into databases available for virtual screening.

Compound databases generally include diverse sets of chemicals. Prior to use in virtual screening applications, these databases are typically filtered to remove compounds with undesirable properties. One filtering approach is to restrict the database to compounds that follow Lipinski's "rule of 5" for "drug-like" compounds [85]. These empirical "rules" reflect chemical properties that were generally common amongst many drugs approved and marketed in the United States in early 2000. Lipinski's "drug-like" properties associated with solubility and permeability included molecular weight ≤ 500 Da., $\log P \leq 5$, hydrogen bond donors ≤ 5 , and hydrogen bond acceptors ≤ 10 . Alternative or additional filtering options include "lead-like" compounds [86-88], vendor source, and calculated indices of adsorption, distribution, metabolic, excretion, toxicity (ADMET) and/or solubility [89, 90]. However, calculated ADMET and solubility indices typically correlate poorly with experimental properties, so these filtering criteria should be used sparingly. Although filtering a database can remove undesirable compounds, such filtering may also eliminate a unique chemical subspace (e.g., all nitro containing compounds) from the database. The reduced chemical space of the filtered database may introduce bias into the virtual screen [91]. Representative members of compounds with undesirable properties (e.g., poorly metabolized functional groups) that occupy unique chemical subspace should be retained in the database during the discovery phase and subsequently modified as part of lead optimization. Filtering criteria must balance the advantage of reduced computational times from screening smaller focused databases with the disadvantage of overlooking potential lead compounds. An appropriate compromise may be achieved by progressively screening less restrictive, more chemically diverse, chemical libraries until suitable numbers of promising lead compounds are generated.

Pharmacophore or fragment databases contain chemical functional groups or very low molecular weight compounds, and are an alternative to chemical libraries for virtual screening applications [92]. Virtual screens with pharmacophore libraries examine a subset of the target protein binding site. In contrast, virtual screens with chemical libraries typically

examine the entire binding site. Novel molecules may be generated with pharmacophore screening by connecting several promising pharmacophore moieties into a single composite molecule that retains the relative orientation of each pharmacophore, [89, 90].

Each molecule within a chemical database must be prepared for the "docking" component of virtual screening, where docking is the automated positioning of a molecule into a defined region of a target biomolecule to maximize a desired pharmacological profile (e.g., binding). Since docking occurs in three-dimensional space, the atoms of each database molecule need defined coordinates which are obtained either experimentally from X-ray crystallography or computationally predicted from atomic connectivity and chemical formula information and programs such as CONCORD [93] or CORINA [94]. Predicted structures are typically minimized with a molecular modeling program to reduce steric interactions and geometric strain. If a docking program treats molecules as rigid entities, then additional rotamer conformations are typically enumerated for each database molecule. These rotomers expand the size of the compound database, and may restrict the number of unique compounds screened. Alternatively, the chemical database must annotate rotatable bonds for use with docking programs that treat molecules as conformationally flexible. Atom "types" and partial atomic charges may need to be included in the compound database depending on the requirements of the docking program's scoring function. Atom "typing" specifies the chemical valence of the atom and partial atomic charge specifies the charge distribution of the molecule. Atom typing and partial charge assignment are typically performed using molecular modeling or docking programs.

DRUG TARGETS

The goal of virtual screening is to discover small molecules that bind to a biomolecule and modulate its function. Proteins typically are targets for activity modulation. High-resolution three-dimensional structures specifying the coordinates of all target protein atoms are necessary for virtual screening. Structures determined by X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy can be obtained experimentally or from public repositories (e.g., RCSB Protein Data Bank, www.rcsb.org/pdb). In X-ray crystal structures the resolution and individual atomic B-factors determine the accuracy of the atomic positions, while NMR structures represent an ensemble of possible conformations. If the targeted region of the protein differs between ensemble conformations, then a range of target structures is used in parallel virtual screens. Bound waters or cofactors are typically removed from the target structure if they do not impact ligand binding or protein stability. In all cases, the target protein coordinates must be augmented with atom type and partial charge information.

Experimentally determined protein structures represent an average of dynamic ensemble conformations. Moreover, many proteins undergo conformational changes upon ligand binding (termed induced fit [95]). Thus, a single rigid structure of the protein target may not accurately reflect the underlying mechanics of protein-ligand binding. A challenge

is to replace rigid protein targets in virtual screens with targets that incorporate conformational dynamics, flexibility, and induced fit. One approach uses molecular dynamics or Monte Carlo simulations to generate an ensemble of different target protein conformations, followed by parallel virtual screens for each protein conformation. More elegant, though not necessarily more accurate, approaches use “soft” docking where repulsion forces between the ligand and protein are reduced [96], normal mode analysis docking [97], or protein side chain flexibly.

DOCKING PROGRAMS

Docking programs automatically predict a preferred orientation (termed “pose”) of a small molecule within a defined site on a target macromolecule and calculate a quantitative measure (termed “score”) for each pose. Poses can be generated using either systematic or random search methods, but determination of the “best” pose is based on maximizing a scoring function that reflects a desired pharmacological profile (e.g., protein-ligand binding). Over the past 20 years, a number of commercial and non-commercial docking programs have been developed; a non-comprehensive listing of current programs include Gold [98], FlexX [99], EUDOC [100, 101], Glide [102], MolDock [103], DOCK [104, 105], Autodock [106, 107], and SODOCK [108].

Each docking program utilizes a unique scoring function (or combination of scoring functions) to rapidly approximate properties such as protein-ligand binding [109]. Common approaches to constructing scoring functions are based on empirical functions and molecular force fields. Empirical scoring functions frequently perform a weighted summing of specific interactions between ligand and target molecules, with weights and interactions derived from defined training sets. These functions can achieve good correlation between calculated scores and measured binding affinities for ligands and protein targets that are similar to those in the training set. However, these scoring functions are limited in their ability to accurately predict binding affinities for diverse sets of molecules and proteins. In contrast, molecular force fields calculate intermolecular “energies” by summing all intermolecular interactions (e.g., van der Waals, electrostatic) that occur between a ligand and protein target. The individual inter-atomic energy terms are parameterized from small molecule test systems and generally independent of a particular protein or ligand. Molecular force field-based scoring functions are dependent upon molecular conformation and inter-atomic distances, and can incorporate approximations for solvation and entropic energy. However, these latter energy terms are poorly understood and difficult to accurately calculate with static pair-wise interactions. Thus, inaccuracies in these energy terms limit the ability of force field-based scoring functions to accurately predict binding affinities. Docking programs have been developed that use combinations of empirical and force field scoring methods [110], and this approach has been reported to perform better than docking programs using a single scoring function [111].

HIGH PERFORMANCE COMPUTING (HPC) RESOURCES FOR DRUG DISCOVERY

Significant computer resources are required to complete virtual screens of large compound databases with exhaustive conformational space searches and complicated scoring functions. Docking simulations using Autodock4 [106, 107] executing on AMD 2.4 GHz Opteron dual core processors require ~1 hr of CPU hr to reproducibly generate poses that agree with published ligand-protein crystal structures [112]. The time required to complete a drug discovery virtual screen scales linearly with the size of the chemical database, thus even less rigorous approaches that complete individual docking simulations within a few minutes require several years of computing time to screen chemical databases containing millions of compounds.

The computational power required to perform large drug discovery virtual screens can be provided by either supercomputers (or large HPC clusters) or distributed computing resources. Modern supercomputers are collections or clusters of hundreds to thousands of interconnected processors. Fast inter-processor connections and shared memory allow processors to perform parallel calculations. Examples of these systems include Beowulf-type clusters of hundreds of identical personal computers and massively integrated systems like IBM’s Blue Gene/P (>65,000 processor cores) supercomputer. In contrast, distributed (or grid) computing utilize geographically separated computers linked through relatively slow Internet connections. A distributed computing resource consists of general-purpose client computers and a designated server. Unused computing cycles are utilized on the client computers for computationally intensive tasks while the server coordinates task scheduling and data exchange with the client computers. IBM’s World Community Grid (www.worldcommunitygrid.org) and BOINC (www.boinc.berkeley.edu) are examples of robust global grid computing environments composed of hundred of thousands of client computers.

Supercomputer and distributed computing resources each provide advantages for drug discovery virtual screening projects. (1) Supercomputers allow efficient inter-processor communication that permit extensive parallel calculations to be performed on interdependent data. This feature can reduce the computational time of force field-based interaction energy calculations. However, virtual screening is inherently a coarse-grained problem where each molecule in the compound database requires independent docking calculations. These highly coarse-grained serial simulations can be efficiently performed by distributed computing systems, and supercomputers can be configured to enable their processors to essentially function as independent serial computers. In both cases, the time required to complete a drug discovery virtual screening project is inversely proportional to the number of available processors, making large grid computing environments competitive with available supercomputers. Individual supercomputer processors are generally dedicated to a single task whereas grid client computers typically execute docking calculations at lower

priority than competing programs and with restricted operating hours. In a defined time interval, a supercomputer processor typically performs more calculations than a grid client processor; our experience suggests that a supercomputer processor completes 5 to 10-fold more calculations in a defined time interval than a grid client computer. (2) Supercomputers provide a uniform operating system and processor environment, which simplify code development, compilation, and optimization. In contrast, a distributed computing environment may contain several different operating systems, requiring program compilation and testing in each operating environment. Additionally, different client computer hardware configurations (e.g., random access memory, free disk space) may prevent calculations from being performed on every distributed computer. (3) Since many client computers within a distributed computing environment communicate over low-bandwidth Internet connections, data transfer between server and client computers must be designed to avoid communication bottlenecks. Comparable bandwidth concerns do not exist with supercomputers. (4) Large HPC clusters are expensive to build, maintain, operate, and upgrade. Thus, large HPC clusters and supercomputers are typically shared among hundred to thousands limited of researchers, which reduced project turnaround time. In contrast, distributed computing systems require inexpensive servers to manage data distribution and access to existing client computers (the clients' primary business functions largely subsidize maintenance, operation, and upgrade of the distributed computer grid hardware).

Clearly, both supercomputer and distributed computing environments support virtual screening calculations. The computing resource used will be dictated by costs and estimates of project completion times given the number of processors that can be accessed by the drug discovery project.

LIMITATIONS OF COMPUTATIONAL DRUG DISCOVERY

The merits of various virtual screening strategies have been addressed in a number of articles [113-117], including a detailed multi-metric comparison of popular docking programs by GlaxoSmithKline researchers [118]. No program was found to significantly and consistently outperform competing programs. Moreover, the relative performance of docking program varied as a function of protein target [118]. The general consensus appears to be that docking programs can predict poses of protein-ligand complexes that largely reproduce co-crystal structures determined by X-ray crystallography. Moreover, virtual screening can produce a reordered database that contains a higher percentage of desired compounds than the original database (i.e., virtual screening enriches the database for molecules with a desired chemical property). Unfortunately, the majority of virtual screening predicted high-affinity protein-binding compounds are false positives, since scoring functions do not accurately predict binding affinities. Extensive experimental testing of predicted inhibitors is required to identify compounds with *in vitro* activity. The high false positive rate of virtual screening is a significant problem when compounds need to be synthesized for experimental testing. In spite of these

shortcomings, there have been some notable successes from using virtual screening to discover new antiviral compounds [119-121].

The shortage of approved antiviral drugs underscores the need to improve drug discovery virtual screening strategies. Generally applicable scoring functions that accurately predict free energies of binding are essential to improve virtual screening [122]. Since force-field-based scoring functions can be applied to diverse bimolecular systems, they are attractive for calculating free energies of binding. Statistical mechanical approaches can calculate the enthalpic and entropic energies associated with ligand, protein, and solvent interactions, and thus provide a methodology to utilize force-field energies for free energy of binding calculations [123].

Accurate statistical mechanical free energies of binding (SM-FEB) calculations require knowledge of the distribution of states accessible to the system. These states can be computationally explored and enumerated using molecular dynamics or Monte Carlo simulations. End point and pathway algorithms are the primary perturbation methods for performing SM-FEB. The end-point method samples a system at the beginning and end of its reaction; the energy difference between the bound and unbound states is correlated with free energy of binding. End point energies can be calculated using linear interaction energy (LIE) [124] or variants of molecular mechanics-Poisson-Boltzmann simulated annealing (MM-PBSA) [125] methods. In contrast, pathway algorithms incrementally perturb the system and calculate the overall energy change along the path between the two end states. These methods can approximate the removal of the atoms that form the binding site as in double annihilation or the decoupling of interactions. For modification of existing compounds, the transformation of one molecule to another can be simulated with chemical alchemy [126], although the modifications are typically limited to single functional groups.

Several concerns have limited widespread use of SM-FEB methods for virtual screening. Although these methods can calculate free energies of binding for poses predicted by docking methods or experimental studies, they have not yet been adapted to perform spatial and conformational searches for optimal ligand-protein poses. Predicted free energies of binding are prone to inaccuracies unless the SM-FEB calculations sufficiently sample the full ensemble of conformational states. These calculations are computationally expensive and require powerful computer resources. Finally, since SM-FEB rely on force-fields to calculate pair-wise interactions, they are limited by the quality of the force-field used.

“DISCOVERING DENGUE DRUGS-TOGETHER” - A FORWARD-LOOKING COMPUTATIONAL PROJECT TO ACCELERATE DENGUE DRUG DISCOVERY

Collaboration between IBM's World Community Grid, the University of Texas Medical Branch, and the University of Chicago established the Discovering Dengue Drugs-Together (DDD-T) distributed computing project (www.worldcommunitygrid.org/DDT) to identify compounds that

inhibit dengue, West Nile (WNV), and hepatitis C (HCV) virus proteases. DDD-T is implemented in two phases that combines extensive virtual screening (Phase 1) and free energy of binding (FEB) calculations (Phase 2) to exploit the strengths of each method. Phase 1 performs virtual screens with each viral protease crystal structure screened against a large database of “commercially available” small molecules. Lowest “energy” scored poses calculated by the docking programs are passed Phase 2 for SM-FEB calculations. This integrated approach has the advantage of using virtual screening as an initial high-throughput filter to provide modest enrichment and poses and SM-FEB calculations as a strict filter to eliminate false positive leads. Proof-of-principle studies using lysozyme and experimentally validated inhibitors have shown that SM-FEB calculations of docked molecules can reliably differentiate between tight, medium, and weak (or non-binding) lysozyme inhibitors (Malmstrom and Watowich, personnel communication). In the current implementation of DDD-T, ~2.2 million compounds are examined in Phase 1 and ~20,000 of the “best” scoring poses are passed to Phase 2. Although some potential small molecules inhibitors may overlooked in Phase 1, sufficiently large numbers of potential inhibitors will be examined to produce numerous promising compounds for experimental validation. To our knowledge, this is the first time a perturbation method has been used to screen thousands of compounds.

DDD-T Phase 1: Virtual Screening Methodology

The docking calculations of Phase 1 are performed using Autodock 4 [106, 107], which employs a genetic algorithm and local minimization to identify optimal poses. Ligands are treated as non-rigid molecules with rotations allowed at non-conjugated dihedral bonds. Polar hydrogen atoms are treated explicitly and non-polar hydrogen atoms are treated with a unified atomic model. The scoring function reflects the “energy” difference between the bound and unbound state of the ligand and is calculated using force-field energies that include van der Waals, hydrogen bonding, and charge-charge (Gasteiger partial charges [127]) interactions and implicit solvation terms. A “penalty” term, determined by the loss of the ligand’s rotational and translational freedom, is included in the scoring function to approximate entropy changes upon ligand binding. Interaction and solvation energies for each atom type are pre-calculated and stored in three-dimensional lookup tables to decrease the time required to compute the scoring function. Cluster analysis is used to group similar orientations into populations since Autodock returns a multitude of predicted poses for each docking simulation; highly populated clusters more accurately represent the correct pose than a single pose with the lowest energy score [128].

The drug targets selected for this project were crystal structures of DEN2V [63, 64, 68], WNV [68, 129] and HCV [130-134] proteases. The project screened a small molecule library that contained 2.2 million drug-like and lead-like molecules extracted from the ZINC database [83]. This library contains predicted atomic coordinates and simple solubility (clogP) predictions for each member compound. In principle, every compound in the library is commercially

available; in our experience a substantial fraction of listed compounds require chemical synthesis.

Docking parameters are adjusted to balance pose accuracy against the time required to complete a single docking simulation. With optimized parameters, each docking simulation requires ~1 hr (benchmarked on 2 GHz 32-bit processors) and generates poses that reproduce >90% of the experimental structures in a trypsin inhibitor test set; longer simulation times do not significantly improve the accuracy of the predicted poses. Each client computer within the computing grid receives a “workunit” for independent processing. Phase 1 workunits contain 5 compounds from the small molecule database and a single target protein for docking; this information is stored and transmitted in kilobyte-sized files. The five docking simulations performed by each workunit typically require less than 8 processing hours on most client computers, although significantly shorter processing times occur on powerful client computers. Completed workunits contain low-scoring poses for each compound and are returned to the parent server in kilobyte-sized files. These small file sizes of pre- and post-processing workunits avoid communication bottlenecks that could result from slow Internet connections. To increase the efficiency of the distributed calculations, internal benchmark control calculations are included within each workunit. These controls reduce the need for redundant workunit calculations to validate the accuracy each client computer, and effectively double the throughput of Phase 1 calculations.

Using ~20% of World Community Grid computational resources, Phase 1 of the DDD-T project can systematically screen a 2.2 million compound library against each protein target in ~10 days. A dedicated HPC cluster with ~10,000 processing units would be required to achieve similar computational throughput. Since the CPU speeds of HPC clusters and grid client computers will experience similar improvements over time, the computational throughput of World Community Grid relative to dedicated HPC clusters will be stable over time.

The poses for each compound’s docking simulation are subjected to cluster analysis (1.5 Å root-mean-square deviation cutoff). From this analysis, the lowest energy pose within the largest cluster is taken as representative of the compound’s bound conformation. The Autodock energy scores of these poses are used to rank-order the compound library and provide an initial enrichment of the library. Twenty thousand compounds (~1% of the original library) with the lowest energy scores are retained for the accurate free energy of binding calculations of Phase 2.

DDD-T Phase 2: SM-FEB Methodology

In Phase 2, thousands of lowest energy compounds from Phase 1 virtual screening are re-evaluated using SM-FEB calculations with conformation restraints to focus the ensemble state sampling [135-137]. Conceptually, these calculations model the binding reaction as two independent processes, namely ligand leaving the binding site and ligand entering the solvent. To reduce the computational effort, explicit inter-atomic energies are calculated for atoms within 12-15 Å of the binding site and boundary potentials are used to estimate more distant interactions. All molecules are

minimized and equilibrated at the start of each process. To model the process of a ligand leaving the protein-binding site or solvent, a decoupling method is used to systematically reduce the interactions between the ligand and the protein/water atoms. The decoupling of charge-charge, dispersion, and repulsion interactions are treated independently. Conformational restraints are also removed within the binding site. Each SM-FEB calculation is divided into dozens of independent decoupling steps, and a molecular dynamics simulation calculated for each step of each process. A weighted histogram method analyzes the molecular dynamic trajectories and calculates the free energy of each decoupling step. The Phase 2 calculations are coarse-grained since each free energy of binding calculation is composed of dozens of independent molecular dynamics simulations, and thousands of independent free energy of binding calculations are required to analyze compounds identified in Phase 1. Thus, Phase 2 of the DDD-T project can be efficiently executed on World Community Grid client computers.

DDD-T Computing Platform

World Community Grid is a philanthropic effort of IBM to provide distributed computing resources to projects designed to benefit humanity. Each day the ~420,000 members (Fig. 8) and >1,000,000 distributed client computers of World Community Grid contribute ~200 years of computer processing time. During Fall 2008 this processing time contributed to five projects on World Community Grid, including "Nutritious Rice for the World", "Help Conquer Cancer", "Human Proteome Folding", "FightAIDS@Home", and "Discovering Dengue Drugs-Together". These coarse-grained projects take advantage of diverse computational methods to perform computer-aided drug discovery, image analysis, protein folding, and genome analysis.

World Community Grid uses a three-tier project management system, with the BOINC program (www.boinc.berkeley.edu) managing communication between the grid server (Tier 2) and client (Tier 3) computers. Tier 1 consists of the servers belonging to the researcher initiating the various projects. In Tier 1, researchers prepare project data for submission to the IBM World Community Grid

servers (Tier 2). For the DDD-T project, ~90 days of CPU time are required to prepare the data for a virtual screen against a single target protein. Tier 2 servers package the pre-processed data from Tier 1 into workunits, distribute the workunits to the client computers (Tier 3) that form the grid, and collect and validate workunit results from client computers. After all results are validated, they are returned for Tier 1 servers for post-processing and final analysis. For the DDD-T project, >200 days of CPU time and ~0.5 TB of disk storage are required to post-process the results from a virtual screen against a single target protein. With IBM scientists controlling Tier 2 servers, the three-tier system provides a secure network for both Tier 1 researchers and Tier 3 clients.

SUMMARY

Although dengue virus is considered one of the most significant mosquito-borne pathogens that threaten global health, there are no approved antiviral drugs to combat dengue infection. Structure-based drug discovery is a promising approach to address this shortcoming given the wealth of structural information available for dengue virus proteins. Unfortunately, most structure-based drug discovery paradigms compromise accuracy of free energy predictions for computational tractability, and thus are largely inefficient and ineffective at unambiguously identifying drug leads. The two-phased approach of the Discovering Dengue Drugs-Together project combines the strengths of virtual screening and molecular dynamic free energy of binding simulations with powerful grid computing resources, to identify antiviral drug leads without extensive false positive predictions. This approach can be easily adapted to combat any pathogen or disease with known virulence factor structures.

ACKNOWLEDGEMENTS

We thank Drs. A. Barrett, J. Briggs, S. Gilbertson, J. Halpert, J. Lee, B. Roux, and J. Stevens for helpful discussions, and the IBM World Community Grid project team (J. Armstrong, V. Bertis, B. Boverman, B. Dolph, T. Hahn, K. Reed, G. Suh, S. Swick, K. Uplinger, N. Wadycki, R. Willner) for supporting the Discovering Dengue Drugs-

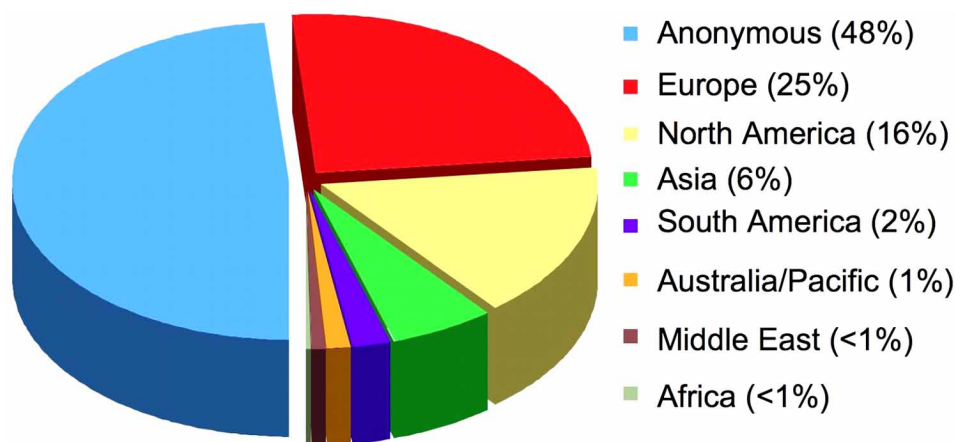


Fig. (8). Relative geographical distribution of client computers registered with World Community Grid in late 2008 (December 2008; www.worldcommunitygrid.org).

Together project. This research was supported in part by a subcontract (SJW) from NIH/NIAID NGA 5 U54 AI057156-03, The Welch Foundation H-1642 (SJW), and training fellowships from the Computational and Structural Biology in Biodefense Training Program (SMT; NIAID T32 AI065396-01) and the Pharmacoinformatics Training Program (RDM; NIH Grant No. T90 DK071504-03) of the W. M. Keck Center for Interdisciplinary Bioscience Training of the Gulf Coast Consortia.

REFERENCES

- Morens, D. M.; Fauci, A. S. *JAMA*, **2008**, 299, 214-216.
- Gubler, D. J.; Meltzer, M. *Adv. Virus Res.*, **1999**, 53, 35-70.
- CDC. Dengue Fever. 2008 24 April 2008 [cited 2008 01 November]; Available from: <http://www.cdc.gov/ncidod/dvbid/dengue>.
- Ligon, B. L. *Semin. Pediatr. Infect. Dis.*, **2005**, 16, 60-65.
- Gubler, D. J.; Clark, G. G. *Emerg. Infect. Dis.*, **1995**, 1, 55-57.
- Gratz, N. G. *Annu. Rev. Entomol.*, **1999**, 44, 51-75.
- Lancioti, R. S.; Roehrig, J. T.; Deubel, V.; Smith, J.; Parker, M.; Steele, K.; Crise, B.; Volpe, K. E.; Crabtree, M. B.; Scherret, J. H.; Hall, R. A.; MacKenzie, J. S.; Cropp, C. B.; Panigrahy, B.; Ostlund, E.; Schmitt, B.; Malkinson, M.; Banet, C.; Weissman, J.; Komar, N.; Savage, H. M.; Stone, W.; McNamara, T.; Gubler, D. J. *Science*, **1999**, 286, 2333-2337.
- Halstead, S. B. *Adv. Virus Res.*, **2003**, 60, 421-467.
- Hayes, E. B.; Gubler, D. J. *Pediatr. Infect. Dis. J.*, **1992**, 11, 311-317.
- Kalayanarooj, S.; Vaughn, D. W.; Nimmannitya, S.; Green, S.; Suntayakorn, S.; Kunentrasai, N.; Viramitrachai, W.; Ratanachuke, S.; Kiatpolpoj, S.; Innis, B. L.; Rothman, A. L.; Nisalak, A.; Ennis, F. A. *J. Infect. Dis.*, **1997**, 176, 313-321.
- Edelman, R. *Clin. Infect. Dis. Suppl.*, **2007**, 1, 56-60.
- McGee, C.; Lewis, M.; Claire, M.; Wagner, W.; Lang, J.; Guy, B.; Tssetsarkin, K.; Higgs, S.; Decelle, T. J. *Infect. Dis.*, **2008**, 197, 693-697.
- Simasathien, S.; Thomas, S.; Watanaverradej, V.; Nisalak, A.; Barberousse, C.; Innis, B.; Sun, W.; Putnak, J.; Eckels, K.; Hutagalung, Y.; Gibbons, R.; Zhang, C.; De La Barrera, R.; Jarman, R.; Chawachalasai, W.; Mammen, M. *Am. J. Trop. Med. Hyg.*, **2008**, 78, 426-433.
- Mukhopadhyay, S.; Kuhn, R. J.; Rossmann, M. G. *Nat. Rev. Microbiol.*, **2005**, 3, 13-22.
- Allison, S. L.; Schalich, J.; Stiasny, K.; Mandl, C. W.; Kunz, C.; Heinz, F. X. *J. Virol.*, **1995**, 69, 695-700.
- Knipe, D.; Howley, P. *Flaviviridae: The Viruses and Their Replication*, in *Fundamental Virology*, 4th ed. Lippincott Williams & Wilkins: Philadelphia, PA, **2001**, pp. 589-639.
- Marks, R. M.; Lu, H.; Sundaresan, R.; Toida, T.; Suzuki, A.; Imanari, T.; Hernaiz, M. J.; Linhardt, R. J. *J. Med. Chem.*, **2001**, 44, 2178-2187.
- Modis, Y.; Ogata, S.; Clements, D.; Harrison, S. C. *Proc. Natl. Acad. Sci. USA*, **2003**, 100, 6986-6991.
- Yang, J. M.; Chen, Y. F.; Tu, Y. Y.; Yen, K. R.; Yang, Y. L. *PLoS ONE*, **2007**, 2, e428.
- Hrobowski, Y. M.; Garry, R. F.; Michael, S. F. *Virol. J.*, **2005**, 2, 49.
- Nowak, T.; Farber, P.; Wengler, G.; Wengler, G. *Virology*, **1989**, 169, 365-376.
- Markoff, L.; Falgout, B.; Chang, A. *Virology*, **1997**, 233, 105-117.
- Kofler, R. M.; Heinz, F. X.; Mandl, C. W. *J. Virol.*, **2002**, 76, 3534-3543.
- Khromykh, A. A.; Westaway, E. G. *J. Virol.*, **1997**, 71, 1497-1505.
- Wang, W. K.; Sung, T. L.; Lee, C. N.; Lin, T. Y.; King, C. C. *Virology*, **2002**, 303, 181-191.
- Ma, L.; Jones, C. T.; Groesch, T. D.; Kuhn, R. J.; Post, C. B. *Proc. Natl. Acad. Sci. USA*, **2004**, 101, 3414-3419.
- Guirakhoo, F.; Bolin, R. A.; Roehrig, J. T. *Virology*, **1992**, 191, 921-931.
- Stadler, K.; Allison, S. L.; Schalich, J.; Heinz, F. X. *J. Virol.*, **1997**, 71, 8475-8481.
- Elshuber, S.; Allison, S. L.; Heinz, F. X.; Mandl, C. W. *J. Gen. Virol.*, **2003**, 84, 183-191.
- Bhardwaj, S.; Holbrook, M.; Shope, R. E.; Barrett, A. D.; Watowich, S. J. *J. Virol.*, **2001**, 75, 4002-4007.
- Rey, F. A.; Heinz, F. X.; Mandl, C.; Kunz, C.; Harrison, S. C. *Nature*, **1995**, 375, 291-298.
- Gromowski, G. D.; Barrett, N. D.; Barrett, A. D. *J. Virol.*, **2008**, 82, 8828-8837.
- Huang, K. C.; Lee, M. C.; Wu, C. W.; Huang, K. J.; Lei, H. Y.; Cheng, J. W. *Proteins*, **2008**, 70, 1116-1119.
- Lok, S. M.; Kostyuchenko, V.; Nybakken, G. E.; Holdaway, H. A.; Battisti, A. J.; Sukupolvi-Petty, S.; Sedlak, D.; Fremont, D. H.; Chipman, P. R.; Roehrig, J. T.; Diamond, M. S.; Kuhn, R. J.; Rossmann, M. G. *Nat. Struct. Mol. Biol.*, **2008**, 15, 312-317.
- Sukupolvi-Petty, S.; Austin, S. K.; Purtha, W. E.; Oliphant, T.; Nybakken, G. E.; Schlesinger, J. J.; Roehrig, J. T.; Gromowski, G. D.; Barrett, A. D.; Fremont, D. H.; Diamond, M. S. *J. Virol.*, **2007**, 81, 12816-12826.
- Perera, R.; Khaliq, M.; Kuhn, R. J. *Antiviral Res.*, **2008**, 80, 11-22.
- Wu, J.; Bera, A. K.; Kuhn, R. J.; Smith, J. L. *J. Virol.*, **2005**, 79, 10268-10277.
- Matusan, A.; Pryor, M.; Davidson, W.; Wright, P. J. *Gen. Virol.*, **2002**, 83, 3093-3102.
- Xu, T.; Sampath, A.; Chao, A.; Wen, D.; Nanao, M.; Chene, P.; Vasudevan, S. G.; Lescar, J. *J. Virol.*, **2005**, 79, 10278-10288.
- Luo, D.; Xu, T.; Hunke, C.; Gruber, G.; Vasudevan, S. G.; Lescar, J. *J. Virol.*, **2008**, 82, 173-183.
- Bartelma, G.; Padmanabhan, R. *Virology*, **2002**, 299, 122-132.
- Sampath, A.; Xu, T.; Chao, A.; Luo, D.; Lescar, J.; Vasudevan, S. G. *J. Virol.*, **2006**, 80, 6686-6690.
- Khromykh, A. A.; Sedlak, P. L.; Westaway, E. G. *J. Virol.*, **1999**, 73, 9247-9255.
- Benarroch, D.; Egloff, M. P.; Mulard, L.; Guerreiro, C.; Romette, J. L.; Canard, B. *J. Biol. Chem.*, **2004**, 279, 35638-35643.
- Egloff, M. P.; Benarroch, D.; Selisko, B.; Romette, J. L.; Canard, B. *EMBO J.*, **2002**, 21, 2757-2768.
- Egloff, M. P.; Decroly, E.; Malet, H.; Selisko, B.; Benarroch, D.; Ferron, F.; Canard, B. *J. Mol. Biol.*, **2007**, 372, 723-736.
- Vasudevan, S. G.; Johansson, M.; Brooks, A. J.; Llewellyn, L. E.; Jans, D. A. *Farmacology*, **2001**, 56, 33-36.
- Yap, T. L.; Xu, T.; Chen, Y. L.; Malet, H.; Egloff, M. P.; Canard, B.; Vasudevan, S. G.; Lescar, J. *J. Virol.*, **2007**, 81, 4753-4765.
- Malet, H.; Masse, N.; Selisko, B.; Romette, J. L.; Alvarez, K.; Guillemot, J. C.; Tolou, H.; Yap, T. L.; Vasudevan, S.; Lescar, J.; Canard, B. *Antiviral Res.*, **2008**, 80, 23-35.
- Bazan, J. F.; Fletterick, R. J. *Virology*, **1989**, 171, 637-639.
- Yusof, R.; Clum, S.; Wetzel, M.; Murthy, H. M.; Padmanabhan, R. *J. Biol. Chem.*, **2000**, 275, 9963-9969.
- Cahour, A.; Falgout, B.; Lai, C. J. *J. Virol.*, **1992**, 66, 1535-1542.
- Chambers, T. J.; Nestorowicz, A.; Amberg, S. M.; Rice, C. M. *J. Virol.*, **1993**, 67, 6797-6807.
- Preugschat, F.; Lenches, E. M.; Strauss, J. H. *J. Virol.*, **1991**, 65, 4749-4758.
- Preugschat, F.; Yao, C. W.; Strauss, J. H. *J. Virol.*, **1990**, 64, 4364-4374.
- Falgout, B.; Pethel, M.; Zhang, Y.; Lai, C. J. *Virology*, **1991**, 65, 2467-2475.
- Hsu, J. T.; Wang, H. C.; Chen, G. W.; Shih, S. R. *Curr. Pharm. Des.*, **2006**, 12, 1301-1314.
- Wlodawer, A.; Vondrasek, J. *Ann. Rev. Biophys. Biomol. Struct.*, **1998**, 27, 249-284.
- Leung, D.; Schroder, K.; White, H.; Fang, N. X.; Stoermer, M. J.; Abbenante, G.; Martin, J. L.; Young, P. R.; Fairlie, D. P. *J. Biol. Chem.*, **2001**, 276, 45762-45771.
- Chanprapaph, S.; Saparpakorn, P.; Sangma, C.; Niyomrattanakit, P.; Hannongbua, S.; Angsuthanasombat, C.; Katzenmeier, G. *Biochem. Biophys. Res. Commun.*, **2005**, 330, 1237-1246.
- Yin, Z.; Patel, S. J.; Wang, W. L.; Wang, G.; Chan, W. L.; Rao, K. R.; Alam, J.; Jeyaraj, D. A.; Ngew, X.; Patel, V.; Beer, D.; Lim, S. P.; Vasudevan, S. G.; Keller, T. H. *Bioorg. Med. Chem. Lett.*, **2006**, 16, 36-39.
- Yin, Z.; Patel, S. J.; Wang, W. L.; Chan, W. L.; Ranga Rao, K. R.; Wang, G.; Ngew, X.; Patel, V.; Beer, D.; Knox, J. E.; Ma, N. L.; Ehrhardt, C.; Lim, S. P.; Vasudevan, S. G.; Keller, T. H. *Bioorg. Med. Chem. Lett.*, **2006**, 16, 40-43.

- [63] Murthy, H. M.; Clum, S.; Padmanabhan, R. *J. Biol. Chem.*, **1999**, *274*, 5573-5580.
- [64] Murthy, H. M.; Judge, K.; DeLucas, L.; Padmanabhan, R. *J. Mol. Biol.*, **2000**, *301*, 759-767.
- [65] Ganesh, V. K.; Muller, N.; Judge, K.; Luan, C. H.; Padmanabhan, R.; Murthy, K. H. *Bioorg. Med. Chem.*, **2005**, *13*, 257-264.
- [66] Watowich, S. personal communication **2009**.
- [67] D'Arcy, A.; Chaillet, M.; Schiering, N.; Villard, F.; Lim, S. P.; Lefeuvre, P.; Erbel, P. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.*, **2006**, *62*, 157-162.
- [68] Erbel, P.; Schiering, N.; D'Arcy, A.; Renatus, M.; Kroemer, M.; Lim, S. P.; Yin, Z.; Keller, T. H.; Vasudevan, S. G.; Hommel, U. *Nat. Struct. Mol. Biol.*, **2006**, *13*, 372-373.
- [69] Tomlinson, S.; Watowich, S. *Biochemistry*, **2008**, *47*, 11763-11770.
- [70] Courageot, M.; Frenkiel, M.; Dos Santos, C.; Deubel, V.; Despres, P. *J. Virology*, **2000**, *74*, 564-572.
- [71] Paranjape, S.; Harris, E. *J. Biol. Chem.*, **2007**, *282*, 30497-30508.
- [72] Talarico, L.; Pujol, C.; Zibetti, R.; Faria, P.; Nosedà, M.; Duarte, M.; Damonte, E. *Antiviral Res.*, **2005**, *66*, 103-110.
- [73] Chu, J.; Yang, P. *Proc. Natl. Acad. Sci. USA*, **2007**, *104*, 3520-3525.
- [74] Kiat, T.; Phippen, R.; Yusof, R.; Ibrahim, H.; Khalid, N.; Rahman, N. *Bioorg. Med. Chem. Lett.*, **2006**, *16*, 3337-3340.
- [75] Laille, M.; Gerald, F.; Debitus, C. *Cell Mol. Life. Sci.*, **1998**, *54*, 167-170.
- [76] Parida, M.; Upadhyay, C.; Pandya, G.; Jana, A. *J. Ethnopharmacol.*, **2002**, *79*, 273-278.
- [77] Reis, S.; Valente, L.; Sampaio, A.; Siani, A.; Gandini, M.; Azeredo, E.; Avila, L.; Mazzei, J. *Int. Immunopharmacol.*, **2008**, *8*, 468-476.
- [78] Hayden, F.; Turner, R.; Gwaltney, J.; Chi-Burris, K.; Gersten, M.; Hsyu, P.; Patick, A.; Smith, G.; Zalman, L. *Antimicrob. Agents Chemother.*, **2003**, *47*, 3907-3916.
- [79] Lamarre, D.; Anderson, P. C.; Bailey, M.; Beaulieu, P.; Bolger, G.; Bonneau, P.; Bos, M.; Cameron, D. R.; Cartier, M.; Cordingley, M. G.; Faucher, A. M.; Goudreau, N.; Kawai, S. H.; Kukolj, G.; Lagace, L.; LaPlante, S. R.; Narjes, H.; Poupert, M. A.; Rancourt, J.; Sentjens, R. E.; St George, R.; Simoneau, B.; Steinmann, G.; Thibeault, D.; Tsantrizos, Y. S.; Weldon, S. M.; Yong, C. L.; Llinas-Brunet, M. *Nature*, **2003**, *426*, 186-189.
- [80] Campbell, S. J.; Gold, N. D.; Jackson, R. M.; Westhead, D. R. *Curr. Opin. Struct. Biol.*, **2003**, *13*, 389-395.
- [81] Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. *Nat. Rev. Drug Discov.*, **2004**, *3*, 935-949.
- [82] Mohan, V.; Gibbs, A. C.; Cummings, M. D.; Jaeger, E. P.; DesJarlais, R. L. *Curr. Pharm. Des.*, **2005**, *11*, 323-333.
- [83] Irwin, J. J.; Shoichet, B. K. *J. Chem. Inf. Model.*, **2005**, *45*, 177-182.
- [84] Allen, F. H. *Acta Crystallogr. B.*, **2002**, *58*, 380-388.
- [85] Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.*, **2001**, *46*, 3-26.
- [86] Oprea, T. I.; Allu, T. K.; Fara, D. C.; Rad, R. F.; Ostropovici, L.; Bologa, C. G. *J. Comput. Aided. Mol. Des.*, **2007**, *21*, 113-119.
- [87] Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D. *J. Chem. Inf. Comput. Sci.*, **2001**, *41*, 1308-1315.
- [88] Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T. *Angew Chem. Int. Ed. Engl.*, **1999**, *38*, 3743-3748.
- [89] van de Waterbeemd, H.; Gifford, E. *Nat. Rev. Drug Discov.*, **2003**, *2*, 192-204.
- [90] Yamashita, F.; Hashida, M. *Drug Metab. Pharmacokinet.*, **2004**, *19*, 327-338.
- [91] Knox, A. J.; Meegan, M. J.; Carta, G.; Lloyd, D. G. *J. Chem. Inf. Model.*, **2005**, *45*, 1908-1919.
- [92] Leach, A. R.; Hann, M. M.; Burrows, J. N.; Griffen, E. J. *Mol. Biosyst.*, **2006**, *2*, 430-446.
- [93] Pearlman, R. *Chem. Des. Auto. News*, **1987**, *2*, 1-6.
- [94] Sadowski, J.; Rudolph, C.; Gasteiger, J. *Tetrahedron Comput. Methodol.*, **1990**, *3*, 537-547.
- [95] Koshland, D. E. *J. Proc. Natl. Acad. Sci. USA*, **1958**, *44*, 98-104.
- [96] Ferrari, A. M.; Wei, B. Q.; Costantino, L.; Shoichet, B. K. *J. Med. Chem.*, **2004**, *47*, 5076-5084.
- [97] Floquet, N.; Marechal, J.; Badet-Denisot, M.; Robert, C.; Dauchez, M.; Perahia, D. *FEBS Lett.*, **2006**, *580*, 5130-5136.
- [98] Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.*, **1997**, *267*, 727-748.
- [99] Kramer, B.; Rarey, M.; Lengauer, T. *Proteins*, **1999**, *37*, 228-241.
- [100] Pang, Y. P.; Perola, E.; Xu, K.; Prendergast, F. G. *J. Comput. Chem.*, **2001**, *22*, 1750-1771.
- [101] Pang, Y. P.; Xu, K.; Kollmeyer, T. M.; Perola, E.; McGrath, W. J.; Green, D. T.; Mangel, W. F. *FEBS Lett.*, **2001**, *502*, 93-97.
- [102] Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.*, **2004**, *47*, 1739-1749.
- [103] Thomsen, R.; Christensen, M. H. *J. Med. Chem.*, **2006**, *49*, 3315-3321.
- [104] Ewing, T. J.; Makino, S.; Skillman, A. G.; Kuntz, I. D. *J. Comput. Aided. Mol. Des.*, **2001**, *15*, 411-428.
- [105] Moustakas, D. T.; Lang, P. T.; Pegg, S.; Pettersen, E.; Kuntz, I. D.; Brooijmans, N.; Rizzo, R. C. *J. Comput. Aided. Mol. Des.*, **2006**, *20*, 601-619.
- [106] Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. *J. Comput. Chem.*, **2007**, *28*, 1145-1152.
- [107] Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.*, **1998**, *19*, 1639-1662.
- [108] Chen, H. M.; Liu, B. F.; Huang, H. L.; Hwang, S. F.; Ho, S. Y. *J. Comput. Chem.*, **2007**, *28*, 612-623.
- [109] Jain, A. N. *Curr. Protein Pept. Sci.*, **2006**, *7*, 407-420.
- [110] Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P. *J. Med. Chem.*, **1999**, *42*, 5100-5109.
- [111] Oda, A.; Tsuchida, K.; Takakura, T.; Yamaotsu, N.; Hirono, S. *J. Chem. Inf. Model.*, **2006**, *46*, 380-391.
- [112] Malmstrom, R.; Watowich, S. **2009**.
- [113] Bursulaya, B. D.; Totrov, M.; Abagyan, R.; Brooks, C. L. 3rd., *J. Comput. Aided. Mol. Des.*, **2003**, *17*, 755-763.
- [114] Chen, H.; Lyne, P. D.; Giordanetto, F.; Lovell, T.; Li, J. *J. Chem. Inf. Model.*, **2006**, *46*, 401-415.
- [115] Joy, S.; Nair, P. S.; Hariharan, R.; Pillai, M. R. *In Silico Biol.*, **2006**, *6*, 601-605.
- [116] Totrov, M.; Abagyan, R. *Curr. Opin. Struct. Biol.*, **2008**, *18*, 178-184.
- [117] Zhou, Z.; Felts, A. K.; Friesner, R. A.; Levy, R. M. *J. Chem. Inf. Model.*, **2007**, *47*, 1599-1608.
- [118] Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. *J. Med. Chem.*, **2006**, *49*, 5912-5931.
- [119] Damm, K.; Ung, P.; Quintero, J.; Gestwicki, J.; Carlson, H. *Biopolymers*, **2008**, *89*, 643-652.
- [120] Luzhkov, V. B.; Selisko, B.; Nordqvist, A.; Peyrane, F.; Decroly, E.; Alvarez, K.; Karlen, A.; Canard, B.; Qvist, J. *Bioorg. Med. Chem.*, **2007**, *15*, 7795-7802.
- [121] Mukherjee, P.; Desai, P.; Ross, L.; White, E.; Avery, M. *Bioorg. Med. Chem. Lett.*, **2008**, *16*, 4138-4149.
- [122] Foloppe, N.; Hubbard, R. *Curr. Med. Chem.*, **2006**, *13*, 3583-3608.
- [123] Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.*, **1997**, *72*, 1047-1069.
- [124] Aqvist, J.; Marelus, J. *Comb. Chem. High Throughput Screen.*, **2001**, *4*, 613-626.
- [125] Thompson, D. C.; Humblet, C.; Joseph-McCarthy, D. *J. Chem. Inf. Model.*, **2008**, *48*, 1081-1091.
- [126] Helms, V.; Wade, R. *J. Am. Chem. Soc.*, **1998**, *120*, 2710-2713.
- [127] Gasteiger, J.; Marsili, M. *Tetrahedron*, **1980**, *36*, 3219-3228.
- [128] Bottegioni, G.; Cavalli, A.; Recanatini, M. *J. Chem. Inf. Model.*, **2006**, *46*, 852-862.
- [129] Aleshin, A. E.; Shiryaev, S. A.; Strongin, A. Y.; Liddington, R. C. *Protein Sci.*, **2007**, *16*, 795-806.
- [130] Arasappan, A.; Njoroge, F. G.; Chen, K. X.; Venkatraman, S.; Parekh, T. N.; Gu, H.; Pichardo, J.; Butkiewicz, N.; Prongay, A.; Madison, V.; Girijavallabhan, V. *Bioorg. Med. Chem. Lett.*, **2006**, *16*, 3960-3965.
- [131] Bogen, S.; Saksena, A. K.; Arasappan, A.; Gu, H.; Njoroge, F. G.; Girijavallabhan, V.; Pichardo, J.; Butkiewicz, N.; Prongay, A.; Madison, V. *Bioorg. Med. Chem. Lett.*, **2005**, *15*, 4515-4519.
- [132] Di Marco, S.; Rizzi, M.; Volpari, C.; Walsh, M. A.; Narjes, F.; Colarusso, S.; De Francesco, R.; Matassa, V. G.; Sollazzo, M. *J. Biol. Chem.*, **2000**, *275*, 7152-7157.
- [133] Kim, J. L.; Morgenstern, K. A.; Lin, C.; Fox, T.; Dwyer, M. D.; Landro, J. A.; Chambers, S. P.; Markland, W.; Lepre, C. A.; O'Malley, E. T.; Harbeson, S. L.; Rice, C. M.; Murcko, M. A.; Caron, P. R.; Thomson, J. A. *Cell*, **1996**, *87*, 343-355.

- [134] Yan, Y.; Li, Y.; Munshi, S.; Sardana, V.; Cole, J. L.; Sardana, M.; Steinkuehler, C.; Tomei, L.; De Francesco, R.; Kuo, L. C.; Chen, Z. *Protein Sci.*, **1998**, *7*, 837-847.
- [135] Deng, Y.; Roux, B. *J. Chem. Phys.*, **2008**, *128*, 115103.
- [136] Wang, J.; Deng, Y.; Roux, B. *Biophys. J.*, **2006**, *91*, 2798-2814.
- [137] Woo, H. J.; Roux, B. *Proc. Natl. Acad. Sci. USA*, **2005**, *102*, 6825-6830.
- [138] Zhang, W.; Chipman, P.; Corver, J.; Hohnson, R.; Zhang, Y.; Suchetana, M.; Baker, T.; Strauss, J.; Rossmann, M.; Kuhn, R. *Nat. Struct. Biol.*, **2003**, *10*, 907-912.
- [139] Zhang, Y.; Zhang, W.; Ogata, S.; Clements, D.; Strauss, J. H.; Baker, T. S.; Kuhn, R. J.; Rossmann, M. G. *Structure*, **2004**, *12*, 1607-1618.
- [140] Modis, Y.; Ogata, S.; Clements, D.; Harrison, S. *Nature*, **2004**, *427*, 313-319.
- [141] Modis, Y.; Ogata, S.; Clements, D.; Harrison, S. *J. Virology*, **2005**, *79*, 1223-1231.
- [142] Pokidysheve, E.; Zhang, Y.; Battisti, A.; Bator-Kelly, C.; Chipman, P.; Ziao, C.; Gregorio, G.; Hendrickson, W.; Kuhn, R.; Rossmann, M. *Cell*, **2006**, *124*, 485-493.
- [143] Kuhn, R.; Zhang, W.; Rossmann, M.; Pletnev, S.; Corver, J.; Lenches, E.; Jones, C.; Mukhopadhyay, S.; Chipman, P.; Strauss, E.; Baker, T.; Strauss, J. *Cell*, **2002**, *108*, 717-725.
- [144] Yu, I.; Zhang, W.; Holdaway, H.; Li, L.; Kostyuchenko, V.; Chipman, P.; Kuhn, R.; Rossmann, M.; Chen, J. *Science*, **2008**, *319*, 1834-1837.
- [145] Li, L.; Lok, S.; Yu, I.; Zhang, Y.; Kuhn, R.; Chen, J.; Rossmann, M. *Science*, **2008**, *319*, 1830-1834.
- [146] Volk, D. E.; Lee, Y. C.; Li, X.; Thiviyanathan, V.; Gromowski, G. D.; Li, L.; Lamb, A. R.; Beasley, D. W.; Barrett, A. D.; Gorenstein, D. G. *Virology*, **2007**, *364*, 147-154.
- [147] Zhang, W.; Corver, J.; Chipman, P.; Zhang, W.; Pletnev, S.; Sedlak, D.; Baker, T.; Strauss, J.; Kuhn, R.; Rossmann, M. *EMBO*, **2003**, *22*, 2604-2613.
- [148] Mueller, N. K.; Pattabiraman, N.; Ansarah-Sobrinho, C.; Viswanathan, P.; Pierson, T. C.; Padmanabhan, R. *Antimicrob. Agents Chemother.*, **2008**, *52*, 3385-3393.